

**ABSTRACT:**

The present invention is a system and method for immersing facial images of people captured automatically from an image or a sequence of images into a live video playback sequence. This method allows viewers to perceive a participation in the viewed "movie" segment. A format is defined for storing the video such that this live playback of the video sequence is possible. A plurality of Computer Vision algorithms in the invention processes a plurality of input image sequences from the means for capturing images, which is pointed at the users in the vicinity of the system and performs the head detection and tracking. The interaction in the invention can be performed either in real-time or off-line depending on the embodiment of the invention in an uncontrolled background.